# Participation in web surveys: a typology

Bosnjak, Michael; Tuten, Tracy L.; Bandilla, Wolfgang

# PARTICIPATION IN WEB SURVEYS A TYPOLOGY[1]

*MICHAEL BOSNJAK, TRACY L. TUTEN & WOLFGANG BANDILLA*

While traditional survey literature has addressed three possible (non-)response patterns (unit nonresponse, item nonresponse, and complete response), Web surveys can capture data about a respondent's answering process, enabling researchers to attain more fine-grained information about individual reactions to such surveys. Based on this information, at least seven (non-)response patterns can be distinguished. This paper describes these seven patterns in a typology of (non-)response. Finally, theoretical and practical implications are discussed.

Während man in der Umfrageforschung im Wesentlichen von drei (Nicht-)Antwortmustern bei Befragungen ausgeht (Unit nonresponse, Item nonresponse und vollständige Befragungsteilnahme), können bei Web-basierten Befragungen (Web surveys) Daten über den Befragungsprozess miterhoben werden, die einen erweiterten Einblick in das tatsächliche Geschehen bei der Beantwortung von Fragen geben. Aufbauend auf diesen Bearbeitungsprozessdaten lassen sich mindestens sieben (Nicht-)Antwortmuster voneinander unterscheiden, die im Rahmen dieses Beitrages beschrieben sowie hinsichtlich ihrer theoretischen und praktischen Implikationen diskutiert werden.

## Introduction

Surveys are generally characterized by the fact that data may be missing for some units of a sample, either partially, or for all variables. This problem of missing data is generally known as 'Nonresponse', whereby one usually differentiates between unit and item nonresponse (Groves/Couper 1998). Unit nonresponse refers to the complete loss of

1 An earlier version of this paper has been presented at the Fifth International Conference on Social Science Methodology in Cologne, Germany, October 4th, 2000. We thank Don A. Dillman and the two anonymous reviewers for valuable comments on an earlier draft of this article. The support of the Center for Survey Research and Methodology (ZUMA) in Mannheim, Germany, during the development and completion of this research and manuscript is gratefully acknowledged.

a survey unit, while item nonresponse refers to missing responses to individual questions. Past work has assumed the existence of three possible responses to requests for survey participation: complete response, item nonresponse, and unit nonresponse. This is the case for all modes of survey data collection, whether it be in-person (e.g., Groves/Cialdini/Couper 1992), by telephone (e.g., Frey 1976; Oksenberg/ Coleman/Cannell 1986), mail (e.g., Armstrong/Overton 1977; Yu/Cooper 1983), or via the Web (e.g., Tuten/Urban/Bosnjak 2001).

With the exception of Web-based surveys, this has been necessarily so, since the process by which a sample member views and answers a self-administered questionnaire has been, for the most part, a black box. However, in Web surveys, the response process can be traced automatically. Such 'para' or 'meta-data' about the answering process can provide insight into the sequencing and completeness of responses. Such data suggest an alternative way of conceptualizing patterns of reacting to a survey which encompasses at least seven possible response types to requests for survey participation. We introduce this typology of response behaviors to more comprehensively describe the potential variations in participation possible in most Web-based surveys. We begin with a brief review of the literature on response behaviors in Web surveys, followed by a description of the response typology. Subsequently, we provide an illustration of the typology proposed. Finally, theoretical as well as practical implications are briefly discussed.

## Response Behaviors in Web Surveys

Findings related to nonresponse in Web surveys are relatively limited. Literature primarily reveals work concerned with explaining (1) volitionally-controlled drop-out behaviors (in this case, the response process is prematurely terminated) and (2) item nonresponse.

Based on a summary of nine Web surveys, Knapp and Heidingsfelder (2001) showed that increased drop-out rates can be expected when using open-ended questions or questions arranged in tables. Dillman et al. (1998) recommended avoiding graphically-complex or ´fancy´ design options. They compared fancy versus plain designs and found higher quit rates when fancy designs were used. This is likely due to the corresponding increase in download time for pages with complex designs.

Dillman (2000) warned of commonly-used techniques in Web surveys which may alienate respondents who are uncomfortable with the Web. The use of pull-down

menus, unclear instructions on how to fill out the questionnaire, and the absence of navigational aids may encourage novice Web-users to break off the survey process.

Frick, Bächtiger and Reips (2001) conducted an experiment on the effect of incentives on response. They concluded that the chance to win prizes in a lottery resulted in lower drop-out rates than in those conditions where no prize draw entry was offered as an incentive. Of particular interest in this context are the opposing findings of an experimental study by Tuten, Bosnjak and Bandilla (2000) which found that the share of unit nonresponders is significantly higher when offering the chance to win a prize, than in cases where altruistic motives for participation are addressed (contribution to scientific research).

Frick, Bächtiger and Reips (2001) also investigated the effect of the order of topics on the amount of drop-outs in a Web survey. In one condition, personal details were requested at the beginning of the investigation (socio-demographic data and e-mail address). In the other condition, these items were positioned at the end of the questionnaire. Surprisingly, drop-outs were significantly lower in the first condition (10.3% versus 17.5%). In other words, when personal data were requested at the beginning, fewer drop-outs occurred.

Bowker and Dillman (2000) conducted an experiment on the effect of the physical layout of answers on the rate of item nonresponse. The results showed a significantly lower item nonresponse rate where response alternatives were right aligned compared to a frequently applied left-aligned format. As a possible explanation for this difference, fewer ´back and forth´ eye movements as well as a decreased number of hand movements associated with completing the questionnaire could have resulted in a lower number of items being missed by some respondents.

A conceptually different approach for the explanation of item nonresponse in Web surveys has been proposed by Bosnjak (2001). In his view, nonresponse in Web surveys is to some extent explainable due to the fact that respondents are not willing to respond to all of the questions presented. As an explanatory approach for this ´noncompliant´ behavior, he uses Ajzen´s theory of planned behavior (Ajzen 1985, Ajzen 1991). The results reported indicate that the theory of planned behavior represents a suitable starting point for predicting the extent of missing data, respectively item nonresponse in open-ended questions, where significant predictions of a medium effect-size could have been demonstrated. In the case of closed-ended questions, however, the theory used seems to have little or no predictive power.

In summary, methods research with regard to Web surveys has focused primarily on detecting some determinants of drop-out and item nonresponse behavior. Certainly Web survey methodology is still in an infancy stage. However, the additional information provided when using the Web to collect data (e.g., automatically-generated log files, visitor tracing programs etc.) can provide a valuable insight into understanding nonresponse and response behaviors. It is no longer necessary to view responses to survey requests within the confines of three conceptually distinct (non)response types. The typology of response behaviors proposed hereafter serves as a descriptive model for operationalizing specific behaviors. This, thereby, ensures increased comparability of individual findings.

## A Typology of Nonresponse Patterns in Web Surveys

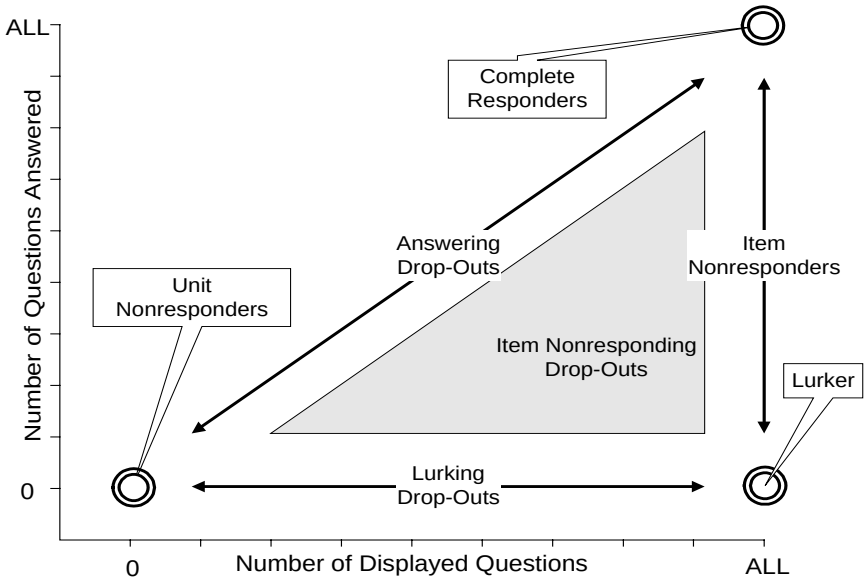
In traditional mail surveys[2], the response process basically remains a mystery. We do not know whether a potential respondent received the questionnaire at all, read it, and began answering it. Such information can hardly be reconstructed afterwards without the aid of a follow-up study. Given this lack of information about the participation process, a survey researcher loses valuable information. If an individual does not return the questionnaire, was it a genuine refusal (i.e., volitionally-controlled) or was it something else (e.g. a technical artifact) to blame? In both cases, the questionnaire is simply categorized as one with unit nonresponse. If a questionnaire is returned incomplete, we do not know whether the participant chose not to answer the remaining questions purposefully, or if he or she merely dropped-out of the process. In either case, the questionnaire is categorized as one with item nonresponse.

One of the substantial advantages of Web surveys, in comparison to mail surveys, is that they can supply para-data, or meta-data, in addition to responses to the substantive questions. There are several methods possible to trace the response process including the use of cgi scripts, java applets, and log files. Regardless of the specific approach used, the data allow the reconstruction of the response process (Batinic/ Bosnjak 1997). In order to log these individual response patterns completely, the following three conditions must be fulfilled: (1) each question must be displayed separately

**2** Due to the commonly accepted notion (see e.g., Bandilla/Bosnjak/Schneid/Stiegler 1999) mail surveys as well as Web surveys belong to the so-called ´self-administered surveys´. Therefore, we are focusing on the possibilities to trace response patterns within this class of surveys and will not elaborate on the possibilities offered by interviewer-administered surveys, like e.g. computer-assisted telephone interviews (CATI) or computer-assisted personal interviews (CAPI).

(screen-by-screen design), (2) the participants are not forced to provide an answer before being allowed to move on (non-restricted design), and (3) each page of the questionnaire must be downloaded separately from the server, and should not be allowed to reside in the Web browser's cache (cache passing pages)[3]. If these conditions are fulfilled, the data set containing information on the user's activities can be used to analyze the completeness and the sequence in which the questions have been processed by the respondents. Figure 1 illustrates the different response patterns.

**Figure 1: (Non-)Response Types in Web Surveys**




In Figure 1, the number of separately displayed questions (abscissa) is set in relation to the number of questions actually answered (ordinate). This graphical representation of observable response patterns allows for a differentiation between the fol-

**3** Various technical implementation methods are available, such as script-based downloading of pages, or integrating specific META tags. The precise technical procedures will not be elaborated upon here.

lowing seven processing types: 1) Complete responders, 2) Unit nonresponders, 3) Answering drop-outs, 4) Lurkers, 5) Lurking drop-outs, 6) Item nonresponders, and 7) Item non-responding drop-outs. Each pattern is described below.

Complete responders are those respondents who view all questions and answer all questions. Unit nonresponders are those individuals who do not participate in the survey. There are two possible variations to the unit nonresponder. Such an individual could be technically-hindered from participation, or he or she may purposefully withdraw after the welcome screen is displayed, but prior to viewing any questions. Answering drop-outs consist of individuals who provide answers to those questions displayed, but quit prior to completing the survey. Lurkers view all of the questions in the survey, but do not answer any of the questions. Lurking drop-outs share some characteristics with ´answering drop-outs´ and ´lurkers´. Such participants view some of the questions without answering, but also quit the survey prior to reaching the end. Item nonresponders view the entire questionnaire, but only answer some of the questions. Item non-responding drop-outs represent a mixture of ´answering drop-outs´ and ´item nonresponders´. Individuals displaying this response behavior view some of the questions, answer some but not all of the questions viewed, and also quit prior to the end of the survey. In our opinion, this typology of response patterns is a more accurate depiction of actual events in Web surveys than the relatively basic categorization of complete participation, unit nonresponse, or item nonresponse.

Using the traditional categorization of possible response behaviors, some behaviors would be mistakenly categorized. Specifically, lurkers and lurking drop-outs would be classified as unit nonresponders. Answering drop-outs and item non-responding drop-outs would be classified the same as item nonresponders. Only ´complete responders´ remain unaffected by the classification system used.

The variations among the seven (non)response types represent significant differences, particularly when one seeks to understand and possibly change response behaviors: unit nonresponders are commonly thought of as people who refused to answer or are hindered from answering. Lurkers and lurking drop-outs, however, are able to respond and are interested enough in the topic to peruse the questions. Yet, they refuse to answer. Lurkers show enough interest to view all questions. Lurking drop-outs either experience technical difficulties in continuing to view the survey or lose interest during the survey, and so do not view all of the questions. Item nonresponders are commonly thought of as people who were not comfortable answering certain questions but otherwise completed the survey. They may have felt a question

was too personal. In other words, we do not tend to assume that item nonresponders lack motivation to respond, but rather that the question(s) influenced their response, or lack thereof. Answering drop-outs, however, begin the survey process much like a complete responder but they drop out prior to completion. These participants may drop-out due to technical difficulties or because they purposefully decide to drop-out. Item non-responding drop-outs begin the survey process like item nonresponders but also quit prior to the end of the survey. This responder type may be more similar to a unit nonresponder than to an item nonresponder.

In all forms of nonresponse and drop-out is always the possibility of both volitional and non-volitional behaviors. With volitionally-controlled, or intentional nonresponse types, the (potential) respondent decides for him or herself to which extent he or she will or will not participate in a survey. Non-volitional nonresponse is caused by technical artifacts, or other external obstacles. For example, an internet session may be disconnected, or a specific net-use pattern may prevent recognition of a request for participation (e.g., Lukawetz 2001). In principle, these two classes of causes must be taken into consideration as an explanation in all drop-out types, as well as for unit nonresponse.

In complete responders, lurkers and item nonresponders, one can assume that all actions are volitionally-controlled due to the evidence that the participants view all questions in the survey. The only possible exception here are the lurkers. While lurkers are not hindered by technical artifacts, it is also possible that lurkers are not always part of the desired sample. Specifically, one must note that it is not necessarily only humans who access, download, and fail to answer the appropriate questionnaire pages, but possibly robots, worms, or wanderers. These non-human, more or less ´intelligent´ (software-based) agents are e.g. used by search engines to find and compile information. Technical procedures for excluding the majority of these robots may exist, yet complete exclusion is not possible due to the diversity of different systems. Dillman (2000) suggested restricting access to people with a valid PIN number in order to reduce coverage error. This technique may be useful for excluding robots.

## An Illustration

A Web-based survey was conducted on the topic of "the roles of men and women in family and work life." The survey questions were arranged according to the design guidelines described above for the identification of different response patterns: (1) each question was displayed separately, (2) participants were not forced to pro-

vide answers before allowed to move on, and (3) each page of the Web questionnaire was protected from being cached. Because our goal was to investigate response patterns, no incentive for participation was offered.

Web users were ‚invited' to the survey through advertising placed on search engines and Web catalogs (e.g., Yahoo, Altavista, etc.). In total, 1469 people participated in the study. Of those answering demographic questions, 35.4% were male and 64.6% female. The mean age in this group was 27.6 years (SD = 8.4 years) and most of the participants were employed (46.5%) or students (34.8%). It is important to note, though, that not all participants are represented in the demographic descriptions. For instance, lurkers viewed the questions, but did not answer them.

Participants were classified into the appropriate segments of the seven-fold typology by analyzing data from both the automatically-generated log file and data set. Specifically, we tracked the questions viewed and answered for each participant. As anticipated, seven specific response types were identifiable.

In this study, 25.3% of the participants were complete responders, 10.2% were unit nonresponders, and 4.3% were answering drop-outs. 6.9% of the respondents were lurkers while 13.3% were lurking drop-outs. 36% of the participants were item nonresponders and 4% were item nonresponding drop-outs (see right column in Table 1).

**Table 1: Comparison of results for the conventional three-fold and the suggested seven-fold typology for the Web survey conducted (n= 1469)**

| **Conventional Three-Fold Typology** | | **Suggested Seven-Fold Typology** | |
|---|---|---|---|
| *(Non)Response Type* | *%* | *(Non)Response Type* | *%* |
| Unit Nonresponse | 30.4 | Unit Nonresponse | 10.2 |
| Item Nonresponse | 44.3 | Item Nonresponse | 36.0 |
| Complete Response | 25.3 | Complete Response | 25.3 |
| | | Lurking Drop-Out | 13.3 |
| | | Lurker | 6.9 |
| | | Answering Drop-Out | 4.3 |
| | | Item Nonresponding Drop-Out | 4.0 |

Using the traditional categories of complete response, unit nonresponse, and item nonresponse (see left column in Table 1), the study described above would have reported unit nonresponse at 30.4% (10.2% unit nonresponders as defined above plus 6.9% lurkers plus 13.3% lurking drop-outs) with a response rate somewhere between 25.3% (according to our definition) and 69.6% (depending upon the degree of unanswered questions in each case). As discussed previously, if using only three response types, lurkers and lurking drop-outs are grouped with unit non-responders. While unit non-responders and lurking drop-outs may have experienced technical difficulties which prevented further participation, it is likely that the three groups differ significantly from each other. If one seeks to minimize nonresponse by encouraging those individuals who are likely to refuse to respond, these differences must be better understood. For instance, given that lurkers do not experience technical problems and willingly choose to view the entire survey, perhaps it is not lack of interest nor motivation, which prevents response but some other attitude.

Similarly, using only item nonresponse, unit nonresponse, and complete response as categories, item nonresponse would have been estimated at 44.3% of returned surveys. Using the seven-fold response typology, we see that 8.3% of the participants answered some questions but dropped out prior to completing the survey. This is an important distinction. The 36% who finished the survey but left missing answers to some questions maintained enough involvement in the survey to complete the activity and did not experience problems completing the survey. However, the answering drop-outs and item nonresponding drop-outs either chose to quit or possibly experienced some problem which interrupted the session. If the drop-out was volitionally controlled, we must learn what variables may have affected that decision. This is especially important for answering drop-outs, as this segment represents individuals who answered all questions up until the decision to quit. Answering drop-outs may be easily converted into complete responders if we develop an understanding of the reasons behind the mechanisms and/or choice to end participation.

## Discussion and Further Suggestions

In conclusion, this paper identifies seven distinct response patterns in Web surveys. The patterns are based upon the questions viewed and answered in a Web survey. In our opinion, the typology suggested here is both of practical and theoretical relevance, as it provides a detailed insight into the individual response patterns in Web surveys, and illuminates the previous 'black box' model of response patterns.

Concerning the explanation of these response patterns, we assume at least two broad complementary approaches to understand and explain nonresponse in Web surveys. The first approach, mostly found in the sociological literature on nonresponse in Web surveys (e.g., Bowker/Dillman, 2000; Dillman, 2000), is primarily focused on design-specific causes of nonresponse error. The second approach, which is in our view complementary to the first, refers closely to the psychological notion of nonresponse as some sort of ´non-compliance´ (see, e.g., Rogelberg/Luong, 1998, for an overview). A combination of these two approaches may give us a comprehensive understanding of the processes underlying the nonresponse phenomenon in general and especially for every single response pattern in Web surveys presented above.

From a practical standpoint, the typology suggested may be used to provide indications of questionnaire quality during the pre-test stage. Changes can then be made, as appropriate, based on the distribution of response types reflected in the pre-test.

## References

Ajzen, I., 1985: From intentions to actions: A theory of planned behavior. pp. 11-39 in: J. Kuhl/J. Beckmann (Eds.), Action control. Berlin: Springer Verlag.

Ajzen, I, 1991: The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50, 179-211.

Armstrong, J. S./Overton, T.S., 1977: Estimating Nonresponse Bias in Mail Surveys. Journal of Marketing Research, 16 (August), 396-402.

Bandilla, B./Bosnjak, M./Schneid, M./Stiegler, A. 1999: Interaktive Medien als Instrument und Gegenstand der empirischen Sozialforschung. [Interactive media as an instrument and a subject-matter for the empirical social sciences] pp. 127-148 in: M. Bergahus (Ed.), Interaktive Medien [Interactive media]. Opladen: Westdeutscher Verlag.

Batinic, B./Bosnjak, M, 1997: Fragebogenuntersuchungen im Internet. [Questionnaire studies on the Internet] pp. 221-244 in: B. Batinic (Ed.), Internet für Psychologen [Internet for Psychologists]. Göttingen: Hogrefe.

Bosnjak, M., 2001: Participation in Non-Restricted Web Surveys: A Typology and Explanatory Model for Item Non-Response. pp. 193-208 in: U.-D. Reips/M. Bosnjak (Eds.), Dimensions of Internet Science. Lengerich: Pabst Science Publishers.

Bowker, D./Dillman, D.A., 2000, May: An Experimental Evaluation of Left and Right Oriented Screens for Web questionnaires. Paper presented at the 55th American Association for Public Opinion Research Annual Conference, Portland, OR.

Dillman, D. A., 2000: Mail and Internet Surveys – The Tailored Design Method. New York, NY: Wiley.

Dillman, D. A./Totora, R.D./Conradt, J./Bowker, D., 1998. Influence of plain versus fancy design on response rates for web surveys. Paper presented at annual meeting of the American Statistical Association, Dallas, TX.

Frey, J., 1976: Survey Research by Telephone. Thousand Oaks, CA: Sage Publications.

Frick, A./Bächtiger, M.T./Reips, U.-D., 2001: Financial Incentives, Personal Information and Drop-Out Rate in Online Studies. pp. 209-220 in: U.-D. Reips/M. Bosnjak (Eds.), Dimensions of Internet Science. Lengerich: Pabst Science Publishers.

Groves, R. M./Cialdini, R.B./Couper, M.P., 1992: Understanding the Decision to Participate in a Survey. Public Opinion Quarterly, 56, 475-495.

Groves, R.M./Couper, M.P., 1998: Nonresponse in Household Interview Surveys. New York: Wiley.

Knapp, F./Heidingsfelder, M., 2001: Drop-out analysis: the effect of research design. pp. 221-230 in: U.-D. Reips/M. Bosnjak (Eds.), Dimensions of Internet Science. Lengerich: Pabst Science Publishers.

Lukawetz, G., 2001: Empirically Quantifying Unit-Nonresponse-Errors in Online Surveys and Suggestions for Computational Correction-Methods. In press in B. Batinic/U.-D. Reips/M. Bosnjak (Eds.), Online Social Sciences. Seattle, WA: Hogrefe & Huber.

Oksenberg, L./Coleman, L./Cannell, C.F., 1986: Interviewers' Voices and refusal rates in telephone surveys. Public Opinion Quarterly, 50 (1), 97-111.

Rogelberg, S.G./Luong, A., 1998: Nonresponse to Mailed Surveys: A Review and Guide. Current Directions in Psychological Science, 7(2), 60-65.

Tuten, T.L./Bosnjak, M./Bandilla, W., 2000: Banner-Advertised Web-Surveys. Marketing Research, 11(4), 17-21.

Tuten, T./Urban, D.J./Bosnjak, M., 2001. Internet Surveys and Data Quality – A Review. In press in: B. Batinic/U.-D. Reips/M. Bosnjak (Eds.), Online Social Sciences. Seattle, WA: Hogrefe & Huber.

Yu, J./Cooper, H., 1983: A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. Journal of Marketing Research, 20 (February), 36-44.